

XAI-FishBD: Explainable Lightweight CNN Framework for Automatic Classification of Indigenous Bangladeshi Small Fish

1st Md. Shihab Shariar
dept. of CSE

Dhaka International University
Dhaka, Bangladesh
shihabshariar543@gmail.com

2nd Md. Soumik Hasan
dept. of CSE

Dhaka International University
Dhaka, Bangladesh
soumik.hasan.smk@gmail.com

3rd Emon Hossain
dept. of CSE

Dhaka International University
Dhaka, Bangladesh
emonhossain4094@gmail.com

4th Md. Shariar Bin Sajjad
dept. of CSE

Dhaka International University
Dhaka, Bangladesh
shariarbinsajjad169@gmail.com

5th Md. Rasel Isalm
dept. of CSE

Dhaka International University
Dhaka, Bangladesh
raselislam07575@gmail.com

6th Md. Abdul Based
dept. of CSE

Dhaka International University
Dhaka, Bangladesh
based@diu.ac

7th Md. Shariful Islam
dept. of CSE

Dhaka International University
Dhaka, Bangladesh
shariful.cse@diu.ac

8th Md. Nahid Hasan
dept. of CSE

Dhaka International University
Dhaka, Bangladesh
nahid.hasan.bondhan@gmail.com

Abstract—The identification of Bangladesh’s native small fish species is essential for preserving biodiversity, maintaining ecological balance, and supporting sustainable aquaculture. The process of manual identification becomes time-consuming because numerous species display identical physical characteristics. The research shows a simple CNN model and explains how it identifies Bangladeshi native small fish species automatically. The framework is optimized for low-latency inference, enabling deployment on edge devices with limited computational capacity. The system uses Gradient-weighted Class Activation Mapping (Grad-CAM) to show the most important areas that affect the model’s classification results. This study used five models, including MobileNetV2, MobileNetV3Small, EfficientNetB0, ShuffleNetV2, and SqueezeNet. The ShuffleNetV2 attains superior results in comparison to other models, achieving 99.82% accuracy together with 0.998255 precision, 0.998241 recall, and a 0.998242 F1-score when tested on a native species dataset. The model uses Grad-CAM to focus on important visual elements, including fin shape and color patterns and body contours. The framework provides an efficient and transparent deep learning pipeline supported by explainability tools.

Index Terms—Automatic Fish Classification, ShuffleNetV2, Grad-CAM, Explainable AI, Indigenous Bangladeshi Fish.

I. INTRODUCTION

Bangladesh contains 260 native freshwater fish species that live in rivers and haor wetlands and beels and floodplains to produce about 4.9 million metric tons of fish during 2022–2023. The Haor wetlands function as essential biodiversity centers, protect 154 species across 35 families, and support aquatic ecosystems along with the communities that depend

on these resources. However, fish populations face growing threats from overfishing, habitat destruction, water pollution, and climate change, endangering species numbers and overall biodiversity [1]–[3].

Bangladesh’s extensive freshwater ecosystem provides high species diversity; however, this ecological context also introduces complexity for automated fish identification. The different environments maintain diverse fish populations that serve as fundamental food sources for people living in these regions. The FAO’s State of World Fisheries and Aquaculture 2020 report shows Bangladesh holds the third position worldwide for open-water fish capture, ranks fifth in aquaculture, fourth in tilapia production, and leads the world in Hilsa production as its national fish. Fish production has grown six-fold over three decades, with 57% from aquaculture, and over 12% of the population relies on fisheries for income and food security [4].

Existing automated fish classification models lack efficient, interpretable frameworks tailored for indigenous small fish, limiting practical field deployment in Bangladesh’s resource-constrained environments. The research goal focuses on building XAI-FishBD, which stands as an explainable lightweight CNN system to perform precise and quick automatic identification of indigenous Bangladeshi small fish species under limited resource field conditions. In this context, automation refers to an end-to-end image-based classification pipeline without manual feature engineering, while lightweight denotes the use of computationally efficient CNN architectures suitable

for deployment in resource-constrained environments.

Most existing studies emphasize commercially important species, leaving a gap in automated classification research for indigenous small freshwater fish. Conventional deep learning architectures are computationally demanding and typically act as black-box models, limiting their usability in field-based low-resource settings. The lack of labeled images for these native species creates a major obstacle, preventing the development of powerful classification systems.

The proposed study presents a novel approach for the automatic classification of indigenous Bangladeshi small fish species using a lightweight and explainable deep learning framework. The key contributions of this work are summarized as follows:

- The research presents a lightweight CNN model that achieves superior performance for identifying small native species through efficient computation.
- The research includes explainable AI methods to make model decisions understandable, leading to better user trust.
- Utilizes two curated datasets of indigenous Bangladeshi small fish, addressing data scarcity, and supports deployment on portable devices for in-situ biodiversity monitoring and conservation efforts.

II. LITERATURE REVIEW

A. Deep Learning-Based Fish Classification

Most fish classification research centers on species that have commercial value, yet scientists have not studied indigenous small fish from Bangladesh enough. Limited datasets and complex, non-interpretable deep learning models hinder accurate and practical classification. Ahmed [5] et al. demonstrated that deep learning models, including DenseNet, Xception, and CNN-LSTM hybrids, achieved up to 97% accuracy in Bangladeshi fish classification. Their study also showed that IoT-enabled systems can support real-time monitoring and operate effectively in practical applications. The *IsVoNet8* CNN model achieved a classification accuracy of 98.62% for eight different marine fish species. Volkan [6] et al. demonstrated that their model surpasses *ResNet50*, *ResNet101*, and *VGG16* through its superior accuracy results and lower loss values, proving that deep learning methods work effectively for fish species classification. Malik et al. [7] developed *FDNet*, which operates as an enhanced *YOLOv7* model to identify nine types of underwater fish species. The system integrates *MobileNetv3* with a BNAM-enhanced *DenseNet* – 169 backbone to achieve better results than *YOLOv3*, *YOLOv4*, *YOLOv5*, *Faster – RCNN*, and *YOLO*.

B. Lightweight CNN Models for Resource-Constrained Environments

Asadullah [8] et al. developed a CNN-based system with *ResNet50*, Logistic Regression, SVC, and CSO, achieving 98.71% accuracy for classifying Bangladeshi freshwater fish.

Transfer learning with pre-trained CNN models *VGG16*, *MobileNet*, and *InceptionV3* was applied to classify betta fish, and Munif et al. [9] showed *InceptionV3* achieved the highest accuracy of 94%. Yamsani et al. [10] developed a CNN-based system that reached 92% accuracy for automated identification of aquatic species. The study demonstrated reliable classification capability but did not evaluate performance under real-time or resource-constrained conditions.

C. Explainable AI in Aquatic Species Classification

Rocha et al. [11] created a Mask R-CNN system to measure freshwater fish body characteristics automatically, reaching 82.21% IOU and 83.89% pixel accuracy. The method effectively handled complex backgrounds with measurement errors under 10%. A hybrid CNN-LSTM model using eye and gill images was developed to classify fish freshness with an 86% accuracy rate, and Biswas et al. [12] showed that integrating CNN spatial features with LSTM temporal learning through interpretable AI methods such as LIME improves food quality assessment. Jayasundara et al. [13] developed FishNET-S and FishNET-T networks for fish quality assessment, achieving 84.1% and 68.3% accuracy, outperforming traditional ML and deep learning models.

D. Research Gap and Motivation

Hossain et al. [14] compared *VGG16*-based classifiers for Bangladeshi fish, achieving 95% accuracy with SoftMax, highlighting trade-offs between accuracy and training time for resource-limited deployment.

Although prior studies report strong classification performance, many models rely on high computational resources and provide limited interpretability. The research introduces a lightweight and explainable CNN system that performs accurate classification of native Bangladeshi fish species in environments with limited resources.

III. METHODOLOGY

A. Proposed Architecture

The overall workflow of the proposed XAI-FishBD framework is illustrated in Fig. 1. The system operates through a single pipeline which combines dataset preprocessing and model training with explainability analysis. Two publicly available datasets—*SmallFishBD* [15] and *BDFreshFish* [16]—were merged to form a comprehensive dataset of 15 indigenous Bangladeshi freshwater fish species. The preprocessing stage preceded the training of five lightweight convolutional neural network (CNN) models which included *MobileNetV2*, *MobileNetV3Small*, *EfficientNetB0*, *ShuffleNetV2* and *SqueezeNet* through transfer learning using ImageNet weights. The models underwent fine-tuning to perform multi-class fish classification tasks which evaluated their predictive accuracy and operational speed and result interpretability. The training process employed the Adam optimizer (learning rate = 1×10^{-3}) and categorical cross-entropy loss function. The evaluation process used accuracy, precision, recall and F1-score metrics to assess performance while Grad-CAM

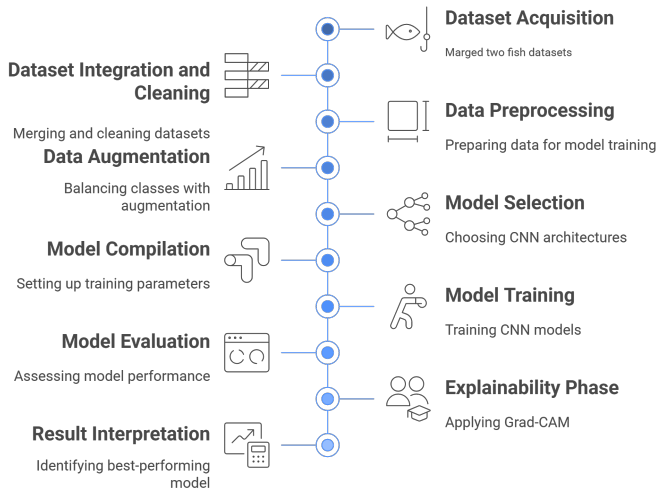


Fig. 1. Overall Workflow Diagram of XAI-FishBD.

visualizations helped to interpret the results. A radar chart later compared overall model performance across metrics.

B. Dataset Description and Preprocessing

The composite data set integrates 15 specific small freshwater fish species documented across Bangladesh's local fish markets and rivers. The classes are denoted as Bele, Chela, Chotka Icha, Guchi, Kachki, Kata Phasa, Koi, Mola, Nama Chanda, Pabda, Puti, Shal Baim, Shing, Taki and Tengra.

Some images recorded at 320×320 px in JPEG format include specific physical features which consist of fin configuration alongside scale pattern and body silhouette. The combined dataset comprises approximately 22,700 images, including both original samples. The datasets were collected from Mendeley Data [15], [16]. The image preprocessing stage included resizing all images to 224×224 px while normalizing their values to the 0,1 interval to match ImageNet-based CNN models. The dataset was partitioned into three sets: 70% for training, 15% for validation, and 15% for testing. Data augmentation was applied only to the training set to address class distribution discrepancies and boost model generalization. The implemented transformations included random horizontal flips, color jitter, random rotations ($\pm 15^\circ$), random zooms and brightness changes to maintain consistent class representation.

Class labels were encoded into 15-dimensional one-hot vectors to align with the multiclass output layer. The CNNs learned discriminative features including fin edges and scale patterns from the data without any manual feature extraction.

C. Applied Model's Architectures

All convolutional neural network (CNN) models underwent training through the same experimental conditions which included $224 \times 224 \times 3$ input dimensions and 32 batch size and 20 training epochs. The same data augmentation and normalization transformations were applied across all architectures. The last classifier layers of each model received

specific modifications to align with the total number of output classes.

The convolution operation common to all architectures is expressed in Equation 1:

$$z_{i,j}^{(k)} = \sum_m \sum_n x_{i+m,j+n} W_{m,n}^{(k)} + b^{(k)} \quad (1)$$

where x denotes the local input region, $W^{(k)}$ is the convolution kernel associated with the k^{th} feature map, and $b^{(k)}$ is the bias term. Spatial downsampling is performed using max-pooling, defined in Equation 2:

$$P_{i,j} = \max_{(m,n) \in R} a_{i+m,j+n} \quad (2)$$

which selects the maximum activation within the receptive region R . The following subsections summarize the applied backbone architectures and their classifier modifications.

1) *MobileNetV2*: MobileNetV2 employs inverted residual blocks and depthwise separable convolutions to achieve computational efficiency. Its feature extraction process can be represented by Equation 3:

$$y = P_{1 \times 1}(D_k(x)) \quad (3)$$

Where $D_k(x)$ denotes depthwise convolution and $P_{1 \times 1}(x)$ denotes pointwise projection. For this study, the original classifier was replaced with a single fully connected layer mapping the extracted features to the target classes.

2) *MobileNetV3-Small*: MobileNetV3-Small integrates squeeze-and-excitation mechanisms and the lightweight hard-swish activation function as shown in Equation 4:

$$h(x) = x \cdot \frac{\text{ReLU6}(x + 3)}{6} \quad (4)$$

This activation improves representational efficiency while preserving low latency. The classification head used in this work consists of a linear layer, ReLU activation, dropout (0.2), and a final linear output layer.

3) *EfficientNet-B0*: EfficientNet-B0 applies the compound scaling rule, which jointly adjusts depth, width, and input resolution according to Equation 5:

$$d = \alpha^\phi, \quad w = \beta^\phi, \quad r = \gamma^\phi \quad (5)$$

with the constraint $\alpha\beta^2\gamma^2 \approx 2$. This coordinated scaling enables favorable accuracy–efficiency characteristics. The classifier was modified by inserting a dropout layer (0.2) followed by a fully connected layer producing class logits.

4) *ShuffleNetV2*: ShuffleNetV2 improves efficiency by channel split operations and channel shuffling, facilitating cross-group information exchange. The shuffle operation is represented as Equation 6:

$$Y = \text{Reshape}(\text{Transpose}(\text{Reshape}(X))) \quad (6)$$

To adapt the architecture to the dataset, the final fully connected layer was replaced with a linear layer corresponding to the number of classes.

5) *SqueezeNet*: SqueezeNet achieves parameter reduction using Fire modules composed of squeeze (1×1) and expand (1×1 and 3×3) convolutional filters. The transformation within a Fire module is expressed in Equation 7:

$$y = f_{1 \times 1}^{\text{exp}}(\sigma(f_{1 \times 1}^{\text{sq}}(x))) + f_{3 \times 3}^{\text{exp}}(\sigma(f_{1 \times 1}^{\text{sq}}(x))) \quad (7)$$

The final classification layer was replaced with a 1×1 convolution producing the required number of output channels, followed by global average pooling.

All five architectures were trained under uniform preprocessing and training configurations. Only the terminal classifier layers were modified, ensuring a controlled and equitable comparison of performance across the selected lightweight CNN models.

IV. RESULTS AND DISCUSSIONS

To evaluate the performance of this framework, five Lightweight CNN models, including MobileNetV2, MobileNetV3Small, EfficientNetB0, ShuffleNetV2, and SqueezeNet were trained and tested on the proposed dataset. The models underwent training through identical experimental setups to achieve an unbiased comparison of their classification results and computational speed and ability to work with new data.

The ShuffleNetV2 model outperformed all other models in every evaluation metric because it achieves the best balance between model complexity and recognition accuracy. The model demonstrates high classification accuracy; however, further evaluation of latency and computational cost is required to confirm suitability for embedded deployment. Table I summarizes the overall performance of all models.

TABLE I
PERFORMANCE COMPARISON OF LIGHTWEIGHT CNN MODELS

Model Name	Accuracy (%)	Precision	Recall	F1-Score
MobileNetV2	99.12%	0.991509	0.991205	0.991262
MobileNetV3Small	99.44%	0.994478	0.994429	0.994431
EfficientNetB0	99.82%	0.998246	0.998241	0.998239
ShuffleNetV2	99.82%	0.998255	0.998241	0.998242
SqueezeNet	93.19%	0.941766	0.931984	0.932234

The ShuffleNetV2 model reached its peak performance with a total accuracy rate of 99.82%, proving its ability to identify closely related small fish species through morphological characteristics. The model's compact design delivers efficiency together with speed and robustness that makes it suitable for deployment in field environments that lack computational resources.

A radar chart serves to display performance comparisons between different metrics through its four axes which represent Accuracy, Precision, Recall and F1-Score. The diagram in Fig. 2 shows the largest enclosed area of ShuffleNetV2, demonstrating the model's ability to maintain balance between all performance metrics. The other models show slight trade-offs when it comes to recall or F1-score.

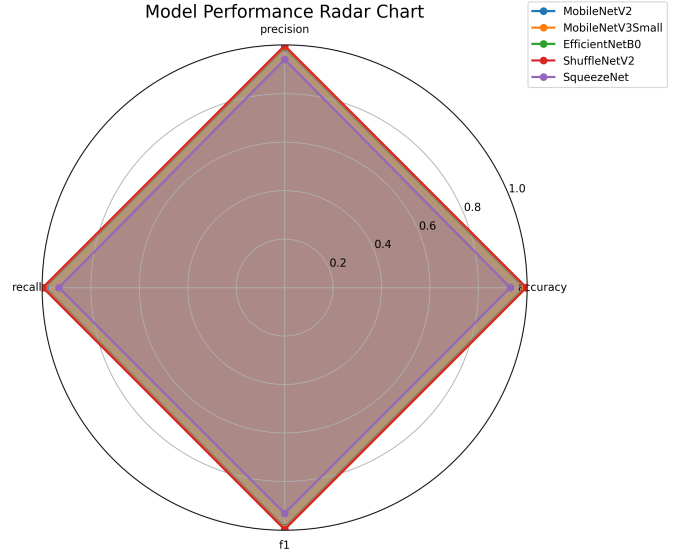


Fig. 2. Radar chart comparing model performances.

ShuffleNetV2 achieves consistently high precision, recall and F1-score, indicating strong generalization across all species classes.

A. Confusion Matrix and ROC Curve of the Best Model

The matrix shows diagonal dominance which demonstrates the model reached almost perfect classification results for every species except for a few errors that occurred between species that share similar physical characteristics including fin texture and color hue variations. To analyze the model's performance in more detail Fig. 3 presents the confusion matrix results for ShuffleNetV2.

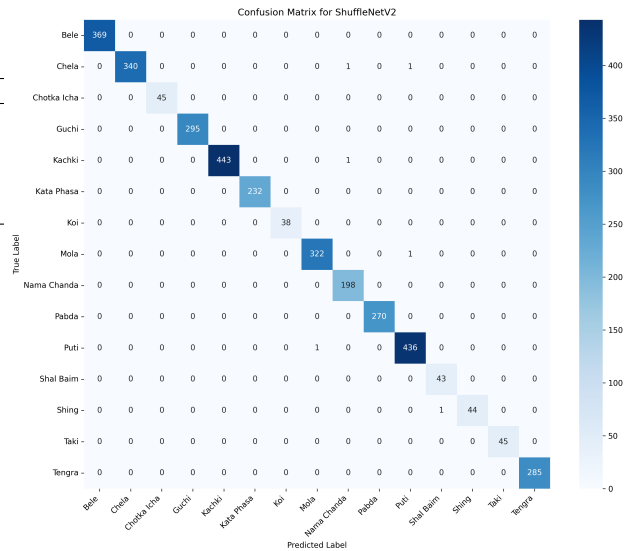


Fig. 3. Confusion matrix of ShuffleNetV2.

The Receiver Operating Characteristic (ROC) curve analysis in Fig. 4 demonstrates that the model achieves superior

discrimination performance. The Area Under the Curve (AUC) value approaches 1.0, confirming that the classifier maintains high sensitivity and specificity across thresholds.

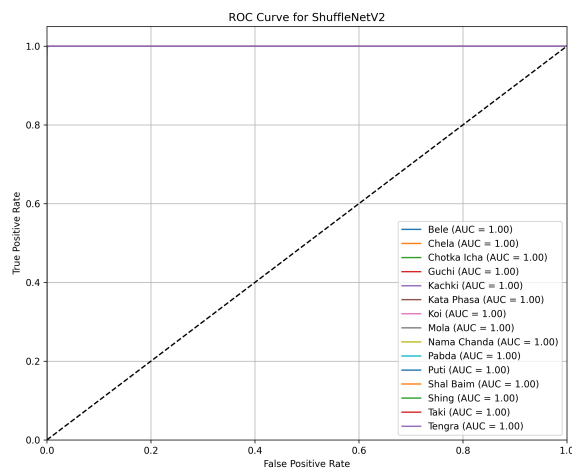


Fig. 4. ROC Curve of ShuffleNetV2.

The evaluation on an independent test set confirms the model's robust performance across all 15 species categories.

B. Training Accuracy and Loss Analysis

The evaluation of training stability appears through Fig. 5 and Fig. 6 which display the training accuracy and training loss curves for the ShuffleNetV2 model. The accuracy curve shows a steady pattern of convergence that moves closer to its maximum point without major fluctuations. The initial epochs show a fast decline in the loss curve which then becomes stable throughout the training process to show successful optimization with no overfitting.

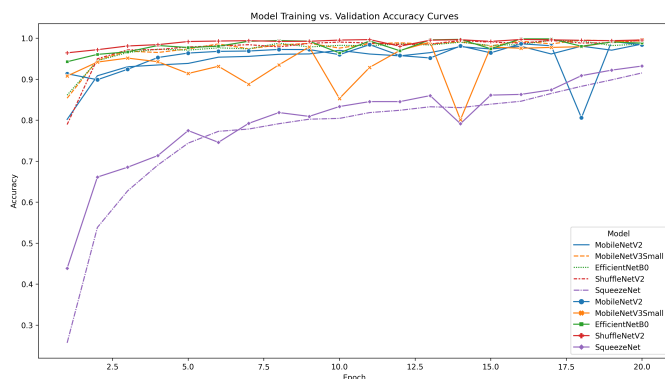


Fig. 5. Training and validation accuracy curve of ShuffleNetV2.

The minor difference between training and validation curves shows the model learned general features which proves its strong performance across different data samples.

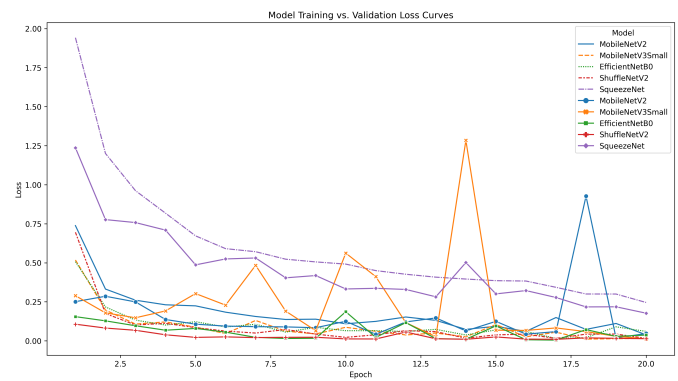


Fig. 6. Training and validation Loss curve of ShuffleNetV2.

C. Explainable AI using Grad-CAM

To understand how the model makes decisions Grad-CAM was used to show the specific areas that influence class predictions. The model learns to identify essential visual elements which include fin shape, body contour and scale texture according to Fig. 7. The highlighted heat regions correspond to commonly used morphological cues; however, expert validation is required to confirm biological relevance.

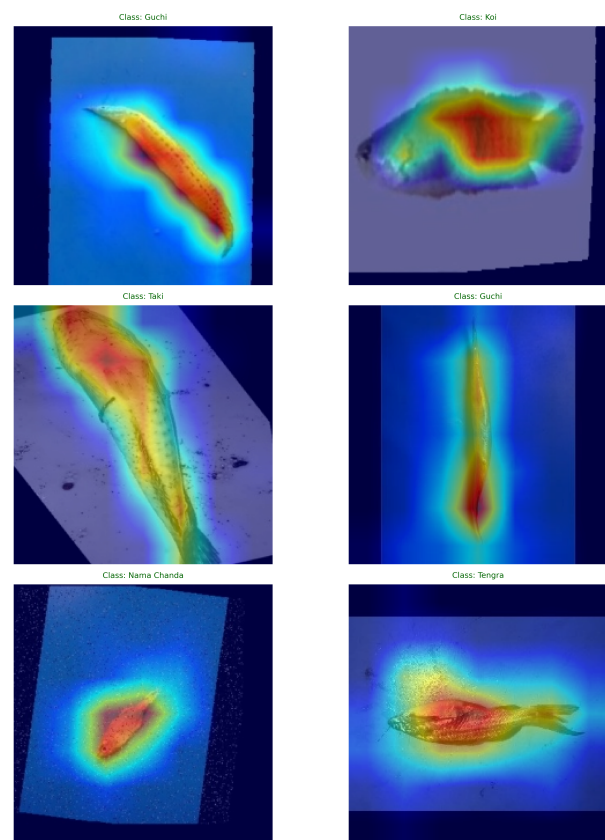


Fig. 7. Grad-CAM visualization.

The Grad-CAM application provides two main benefits because it verifies the model's decision-making process and

offers essential explanations for scientific and ecological work which demands open transparency. The XAI-FishBD system functions as a reliable AI tool which surpasses basic black-box classification to support biodiversity research.

D. Comparison with Existing Works

While many prior models achieve strong accuracy, most do not incorporate explainability frameworks suitable for field deployment. The XAI-FishBD framework delivers an efficient lightweight system that provides explainable results and high performance for Bangladeshi native small fish.

TABLE II
COMPARISON WITH EXISTING FISH CLASSIFICATION STUDIES

Study / Year	Model Used	Dataset Type	Explainability	Accuracy (%)
Yamsani et al., 2025 [10]	CNN Based Models	Marine and freshwater	No	92%
Hossain et al., 2023 [14]	SoftMax Classifier	Custom dataset, (Bangladesh)	No	95%
Hasanul et al., 2024 [17]	EfficientNetV2B0	SmallFishBD [15]	No	98.43%
Wahidur et al., 2024 [5]	CNN + LSTM	BDFreshFish [16]	No	97%
Proposed XAI-FishBD, 2025	ShuffleNetV2, Grad-CAM	Merged two Native small fish dataset (Bangladesh) [15], [16]	Yes	99.82%

The proposed XAI-FishBD framework demonstrates better performance than all existing methods based on accuracy and efficiency, interpretability results as shown in Table II. The model achieves 99.82% accuracy with ShuffleNetV2 lightweight architecture by merging two Bangladeshi small-fish datasets which enables real-time applications that need low resource consumption. The Grad-CAM integration shows biologically meaningful regions that proves the model's reliability through its transparent explanation. The integration of lightweight CNNs with Grad-CAM allows XAI-FishBD to balance accuracy with interpretability.

V. CONCLUSION

This paper presented XAI-FishBD, an explainable and lightweight CNN-based framework for automatic classification of 15 indigenous Bangladeshi small fish species. The research evaluated five lightweight CNN models MobileNetV2, MobileNetV3Small, EfficientNetB0, ShuffleNetV2 and SqueezeNet using the same experimental setup with a combined native fish dataset.

Experimental results demonstrate that ShuffleNetV2 achieved the best overall performance, reaching an accuracy of 99.82%, with precision, recall, and F1-score values of 0.9983, 0.9982 and 0.9982 respectively. The experimental results show that the developed system can identify small fish species which look alike while operating at speeds that enable its use in resource-limited settings. The model gains improved transparency through Grad-CAM integration which shows essential biological features including fin structure and body contours to help users understand the model better and trust its results.

The evaluation results show favorable outcomes; however, further testing must be conducted to assess the model performance in practical real-world applications. The upcoming

research will investigate cross-validation methods together with domain-shift evaluation and external dataset testing and optimization for deployment which includes latency assessment and model compression techniques. The research will enhance XAI-FishBD system performance for biodiversity monitoring and sustainable fisheries management.

REFERENCES

- [1] S. Nahar, M. H. Rahman, M. S. Islam, M. A. Alam, S. T. Lupa, A. Amin *et al.*, "Fish biodiversity at haor wetland of bangladesh: A comprehensive review," *International Journal of Aquaculture and Fish Science*, vol. 11, no. 1, pp. 001–009, 2025.
- [2] F. J. Abdullah-Al-Hasan, A. Azim, K. Mukherjee, M. M. Rahman, M. J. Hossain, A. Saha, M. H. Rahman, and M. H. Islam, "Assessment of fisheries biodiversity and associated marketing systems in the bishkhali river ecosystem of coastal bangladesh," *International Journal of Fish and Aquatic Studies*, vol. 13, no. 3, pp. 123–130, 2025.
- [3] WorldFish, "Bangladesh," Online, 2025, accessed: Oct. 24, 2025. [Online]. Available: <https://worldfishcenter.org/where-we-work/asia/bangladesh>
- [4] S. Hossain, "Fisheries industry in bangladesh: Current trends and future opportunities," Online, 2023, accessed: Oct. 24, 2025. [Online]. Available: <https://businessinspection.com.bd/fisheries-industry-in-bangladesh/>
- [5] M. A. Ahmed, M. S. Hossain, W. Rahman, A. H. Uddin, and M. T. Islam, "An advanced bangladeshi local fish classification system based on the combination of deep learning and the internet of things," *Journal of Agriculture and Food Research*, vol. 14, p. 100663, 2023.
- [6] V. Kaya, İ. Akgül, and Ö. Z. Tanır, "İsvonet8: A proposed deep learning model for classification of some fish species," *Journal of Agricultural Sciences*, vol. 29, no. 1, pp. 298–307, 2023.
- [7] H. Malik, A. Naem, S. Hassan, F. Ali, R. A. Naqvi, and D. K. Yon, "Multi-classification deep neural networks for identification of fish species using camera captured images," *PLoS ONE*, vol. 18, no. 4, p. e0284992, 2023.
- [8] A. Shaikh, W. Rahman, K. Roksana, T. Islam, M. M. Rahman, H. Al-shahrani, and M. S. A. Reshan, "An improved deep cnn-based freshwater fish classification with cascaded bio-inspired networks," *Automatika*, vol. 66, no. 2, pp. 249–280, 2025.
- [9] R. Munif and A. Prahara, "Betta fish classification using transfer learning and fine-tuning of cnn models," *Science in Information Technology Letters*, vol. 5, no. 1, pp. 16–25, 2024.
- [10] N. Yamsani, P. Nasra, and S. Gupta, "Automated fish species classification using convolutional neural networks," in *Proc. 3rd Int. Conf. Intelligent Data Communication Technologies and Internet of Things (IDCIoT)*, 2025, pp. 1653–1658.
- [11] W. S. Rocha, T. F. C. da Fonseca, C. Y. V. Watanabe *et al.*, "Automatic measurement of fish from images using convolutional neural networks," *Multimedia Tools and Applications*, vol. 84, pp. 6327–6347, 2025.
- [12] S. Biswas, C. Nahata, S. Ghosh, S. Sahoo, and D. Biswas, "Fish freshness detection via hybrid cnn-lstm: An interpretable deep learning model," in *Biologically Inspired Techniques in Many Criteria Decision-Making*, ser. Learning and Analytics in Intelligent Systems. Springer, 2025, vol. 45.
- [13] J. M. V. D. B. Jayasundara, R. M. L. S. Ramanayake, H. M. N. B. Senarath *et al.*, "Deep learning for automated fish grading," *Journal of Agriculture and Food Research*, vol. 14, p. 100711, 2023.
- [14] S. Hossain, K. Shahadat, and M. S. Rahman, "Deep learning-based local fish classification: A comparative study of vgg16 models and multiple classifiers," in *Proc. IEEE Region 10 Humanitarian Technology Conf. (R10-HTC)*, 2023, pp. 589–594.
- [15] M. H. Ferdous, R. H. Prito, M. Ahmed *et al.*, "Smallfishbd: A comprehensive image dataset of common small fish varieties in bangladesh," Mendeley Data, v2, 2024.
- [16] M. W. Rahman, M. T. Islam, R. Khan, and M. M. Rahman, "Bdfreshfish: A comprehensive image dataset for machine learning applications," Mendeley Data, v1, 2024.
- [17] M. H. Ferdous, R. H. Prito, M. Ahmed *et al.*, "Smallfishbd: An extensive image dataset of common native small fish species in bangladesh," *Data in Brief*, vol. 63, p. 112193, 2025.